

SIZE CONSISTENCY AND THE CI METHOD.§1 The size consistency problem

A condition to be fulfilled by any method in quantum chemistry is the following : the quality of the results should be independent of the size of the system, that is, the number of nuclei and electrons in the molecule. This condition is not automatically satisfied, as we will see below. In general it is not easy to check whether the condition is satisfied in practice. However, for some model systems, it may be checked theoretically.

We will discuss a model system containing two subsystems, which do not interact (which are far from each other). The total energy E_{AB} should then satisfy

$$E_{AB} = E_A + E_B \quad (1.1)$$

This condition is called the size consistency condition.

If the subsystems are far from each other, the Hamiltonian may be taken as :

$$\mathbf{H}_{AB} = \mathbf{H}_A(1\dots n_A) + \mathbf{H}_B(n_A + 1\dots n_A + n_B) \quad (1.2)$$

Eq (1.2) implies that the electrons may be divided into two groups, each belonging to one of the subsystems, and that the interactions in \mathbf{H}_{AB} between the particles of one subsystem with the particles of the other subsystem may be ignored. Moreover, exchange effects between the two subsystems may then be ignored. Therefore the normalised wave function for the total system may be represented by :

$$\Psi_{AB} = \Psi_A(1\dots n_A)\Psi_B(n_A + 1\dots n_A + n_B) \quad (1.3)$$

if Ψ_A and Ψ_B are also normalised and if

$$\begin{aligned} \mathbf{H}_A(1\dots n_A)\Psi_A(1\dots n_A) &= E_A\Psi_A(1\dots n_A) \\ \mathbf{H}_B(n_A + 1\dots n_A + n_B)\Psi_B(n_A + 1\dots n_A + n_B) &= E_B\Psi_B(n_A + 1\dots n_A + n_B) \end{aligned} \quad (1.4)$$

where it is assumed that the orbitals in Ψ_A and Ψ_B are localised on either subsystem. From Eqs (1.2) and (1.3) it then follows that Eq (1.1) holds.

Here we have used the exact eigenfunctions for the subsystems. However, Eq (1.1) may be satisfied even if Eq (1.4) does not hold, provided that Eq (1.3) holds. For variational methods the energy is calculated as the expectation value of the optimized wave function and Eqs (1.2) with (1.3) then imply Eq (1.1) :

$$E_{AB} = \langle \Psi_A \Psi_B | \mathbf{H}_A + \mathbf{H}_B | \Psi_A \Psi_B \rangle = \langle \Psi_A | \mathbf{H}_A | \Psi_A \rangle + \langle \Psi_B | \mathbf{H}_B | \Psi_B \rangle = E_A + E_B \quad (1.5)$$

Both Eqs (1.3) and (1.5) may be used in order to check whether a variational method is size consistent.

Size Consistency in CI

As will be seen below, the two extremes, the HF method and the Full CI (FCI) method, are size consistent. However, any CI with the same restriction on the excitation level for A, B and A+B is not.

§2 The HF method

The HF method satisfies the size consistency condition. In order to show this we start with orbitals localised on the subsystems. First it is to be shown that Eq (1.3) may be used.

The HF function for the total system is given by :

$$\Psi_{AB}^{\text{HF}} = |\dots \psi_i^A \dots \psi_j^B \dots| \quad (1.6)$$

The antisymmetrisation generates exchange integral contributions to the HF energy, otherwise the energy is the same as for a simple product function. Eq (1.6) differs from the product form of Eq (1.3) by terms containing permutations of electrons on A with electrons on B. However, these permuted terms in Eq (1.6) do not contribute to the energy, since the subsystems are far apart. Therefore the product form of Eq (1.3) yields the same energy as the function of Eq (1.6) and thus Eq (1.5) is satisfied.

If delocalised MSOs are used, the complete function Eq (1.6) has to be used. However, the determinant is invariant under transformations of the occupied orbitals. Therefore the orbitals may be transformed back to the localised ones and the HF determinant of the total system will still yield the same energy, satisfying Eq (1.5).

Note that the HF method is variational, that is, the energy is minimised for each case. Since the variational space for the total system is the sum of the subsystem variational spaces, the HF MOs for AB will be equivalent to the direct sum of the monomer HF MOs.

§3 The FCI method

The FCI method gives the best possible result for a given basis set (basis set limit). As expected for a method which is in principle exact (for an infinite basis set), this method also satisfies Eq (1.5). This may be rationalised as follows. For the total system the same MOs are used as for the subsystems. The FCI dimer configuration set contains all possible configurations with N_A electrons in orbitals on subsystem A and with N_B electrons in orbitals on subsystem B. This is the direct product of the variational spaces of the subsystems. As in the HF case each configuration may be replaced by the corresponding product of functions of the form of Eq (1.3). There are also configurations corresponding to A^+B^- etc. However, these do not interact with the previous ones, since the subsystems are far apart. Therefore Eq (1.3) may be used if localised MOs are used in the dimer calculation. Moreover, for a given AO basis set the FCI result is

Size Consistency in CI

independent of the MOs. Therefore Eq (1.5) also holds if delocalised MOs are used in the dimer calculation.

Exercise 1.1

Show that Eq (1.3) satisfies Eq (1.1) if Eq (1.4) holds.

§4 The DCI method

A CI function is a linear combination of determinants, where the coefficients are determined variationally (by minimizing the energy).

$$\Psi_{\text{CI}} = \sum_K c_K \Phi_K \quad (1.7)$$

with

$$\Phi_K = |\psi_1 \dots \psi_1 \dots \psi_N| \quad (1.8)$$

where the ψ_i are the Molecular Spin Orbitals (MSOs). We will now discuss the Double excitation CI (DCI) method. The corresponding wave function may be written as :

$$\Psi_{\text{DCI}} = c_0 \Phi_0 + \sum_{I,J} \sum_{A,B} c_{I,J}^{A,B} \Phi_{I,J}^{A,B} \quad (1.9)$$

where Φ_0 and Φ_{IJ}^{AB} correspond to the HF function and the doubly excited determinants respectively.

The essential disadvantage of such CI calculations is, that the degree of excitation w.r.t. the HF determinant is constrained. This constraint makes for tractable calculations, but the variational treatment then destroys the size consistency.

This may be shown by a simple example : two H_2 molecules without interaction, using the minimal AO basis set. For each molecule two determinants are needed :

$$\Phi_g = |\sigma_g \overline{\sigma_g}| \quad (1.10)$$

$$\Phi_u = |\sigma_u \overline{\sigma_u}|$$

where

$$\sigma_g = N_g (s_A + s_B) \quad (1.11)$$

$$\sigma_u = N_u (s_A - s_B)$$

Φ_g is the HF determinant, Φ_u is the only double excitation and the single excitation is not needed, since it has a different symmetry.

The monomer FCI result is obtained by solving the 2×2 secular problem for the two functions in Eq (1.10). The lowest eigenvalue is denoted by E_f . The corresponding (monomer) wave function is given by :

Size Consistency in CI

$$\Psi_f = c_g \Phi_g + c_u \Phi_u \quad (1.12)$$

Since the molecules do not interact, the total FCI energy of the two molecules is then given by :

$$E_F = 2E_f \quad (1.13)$$

However, the DCI energy for the total system turns out to be higher than E_F .

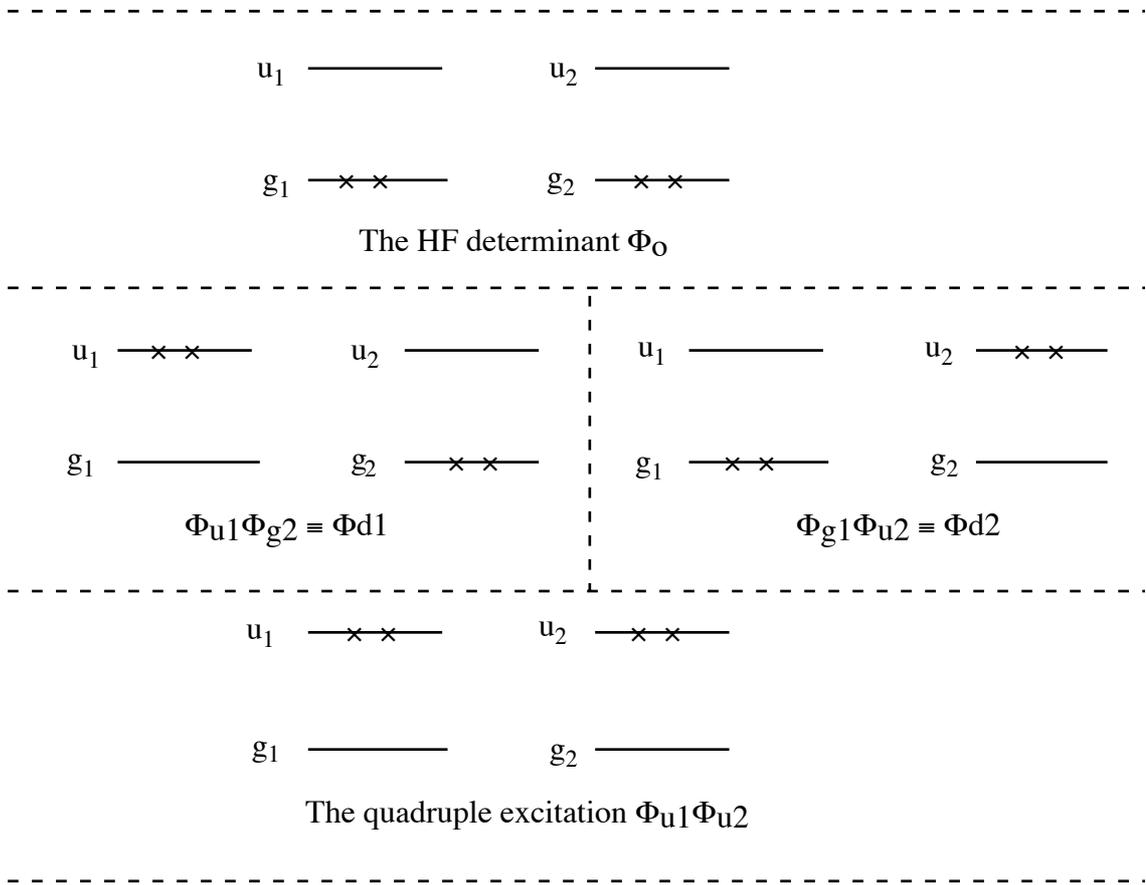


Fig 2.1 The determinants in the FCI for two H₂ molecules.

The correct wave function Ψ_F for the dimer is given by :

$$\begin{aligned} \Psi_F &= \Psi_{f1} \Psi_{f2} = [c_g \Phi_{g1} + c_u \Phi_{u1}] [c_g \Phi_{g2} + c_u \Phi_{u2}] = \\ &= c_g^2 \Phi_{g1} \Phi_{g2} + c_g c_u [\Phi_{u1} \Phi_{g2} + \Phi_{g1} \Phi_{u2}] + c_u^2 \Phi_{u1} \Phi_{u2} \end{aligned} \quad (1.14)$$

whereas the DCI wave function corresponds to :

$$\Psi_{DCI} = c_0 \Phi_0 + c_d (\Phi_{d1} + \Phi_{d2}) \quad (1.15)$$

The HF function Φ_0 is identical to $\Phi_{g1}\Phi_{g2}$. The double excitation on molecule 1, Φ_{d1} is identical to $\Phi_{u1}\Phi_{g2}$ and Φ_{d2} is identical to $\Phi_{g1}\Phi_{u2}$. The coefficients for Φ_{d1} and Φ_{d2} are equal,

Size Consistency in CI

since the molecules are equivalent. The remaining function $\Phi_{u1}\Phi_{u2}$ in Eq (1.14) does not appear in Eq (1.15), since it is a quadruple excitation. Therefore the DCI function does not yield the correct energy. In fact the DCI energy is higher than the FCI energy since one configuration is missing. Also, the coefficient c_d in Eq (1.15) will be different from the coefficient $c_g c_u$ in Eq (1.14). However, there is no reason why the number of independent variables in the dimer calculation should be higher than in the monomer calculation, since we know a priori that the dimer energy equals twice the monomer energy. First we note that the monomer calculation has two coefficients, but that the normalisation condition determines a relation between these two. Therefore we have only 1 independent variable. Because of Eq (1.14) the FCI dimer calculation also has 1 independent variable, just like the monomer calculation. The coefficient for the quadruple excitation $\Phi_{u1}\Phi_{u2}$ may be calculated directly from the other coefficients :

$$c_u^2 = (c_u c_g)^2 / c_g^2 \quad (1.16)$$

Therefore Ψ_F contains as much information as either Ψ_f or Ψ_{DCI} . This fact may be used in order to calculate a correction to the DCI energy.

Note that the size consistency problem is inherent to the CI method if the degree of excitation is somehow restricted, since the function for the total system will always contain higher excitations than were present in the monomer calculations.

§.5 CI calculations with intermediate normalisation

The norm of the wave function must be finite. The usual normalisation to 1 is useful for interpretational purposes and for variational calculations using the Lagrange multiplier method. However, it is not always the most practical choice. In the methods designed to yield size consistent results usually intermediate normalisation is used, as in perturbation theory. The intermediate normalisation is defined by :

$$\langle \Psi | \Phi_0 \rangle = 1 \quad (1.17)$$

where Φ_0 is the reference function (HF in Single Reference CI). The relation to the usual normalisation is given by :

$$|\Psi_{var}\rangle = c_0 |\Phi_0\rangle + \sum_k c_k |\Phi_k\rangle = c_0 [|\Phi_0\rangle + \sum_k d_k |\Phi_k\rangle] = c_0 |\Psi_{IN}\rangle$$

with

$$d_k = \frac{c_k}{c_0} \quad (1.18)$$

or

$$c_0 = \langle \Psi_{IN} | \Psi_{IN} \rangle^{-1/2} = (1 + \sum_k d_k^2)^{-1/2}$$

Size Consistency in CI

where Ψ_{var} is normalised to 1 and Ψ_{IN} has intermediate normalisation. The advantage of the intermediate normalisation is that it simplifies the construction of a size consistent wave function.

In order to simplify the comparison between CI functions and size consistent functions we now rewrite the CI secular equations to the form which corresponds to the intermediately normalised CI function Ψ_{IN} with $E_0 = \langle \Phi_0 | \mathbf{H} | \Phi_0 \rangle$ as the reference energy. The CI equations in matrix form are given by :

$$\begin{bmatrix} E_0 & \mathbf{H}_{\text{int}}^{\leq} \\ \mathbf{H}_{\text{int}} & \mathbf{H}_{\text{ex}} + E_0 \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{c}_0 \\ \mathbf{c} \end{bmatrix} = E_{\text{CI}} \begin{bmatrix} \mathbf{c}_0 \\ \mathbf{c} \end{bmatrix}$$

with (1.19)

$$(\mathbf{H}_{\text{int}})_{kl} = \langle \Phi_k | \mathbf{H} | \Phi_l \rangle$$

$$(\mathbf{H}_{\text{ex}})_{kl} = \langle \Phi_k | \mathbf{H} - E_0 | \Phi_l \rangle = \langle \Phi_k | \mathbf{H} | \Phi_l \rangle - E_0 \delta_{kl}$$

E_0 is the HF energy and \mathbf{I} is the unit matrix.

Eq (1.19) is equivalent to :

$$\begin{bmatrix} 0 & \mathbf{H}_{\text{int}}^{\leq} \\ \mathbf{H}_{\text{int}} & \mathbf{H}_{\text{ex}} \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{d} \end{bmatrix} = E_{\text{corr}} \begin{bmatrix} 1 \\ \mathbf{d} \end{bmatrix}$$

with (1.20)

$$E_{\text{corr}} = E_{\text{CI}} - E_0$$

$$\mathbf{d}_k = \frac{\mathbf{c}_k}{\mathbf{c}_0}$$

The solution of Eq (1.20) is equivalent to the solution of Eq (1.19)

Exercise 1.2

Derive Eq (1.20) from Eq (1.19)

Eq (1.20) suggests an alternative way of finding the solution for the CI problem. Eq (1.20) may be separated into :

$$\mathbf{H}_{\text{int}}^{\leq} \mathbf{d} = E_{\text{corr}} \mathbf{d} \tag{1.21a}$$

$$\mathbf{H}_{\text{int}} + \mathbf{H}_{\text{ex}} \mathbf{d} = E_{\text{corr}} \mathbf{d} \tag{1.21b}$$

or equivalently :

Size Consistency in CI

$$\mathbf{d} = (\mathbf{E}_{\text{corr}} \mathbf{I} - \mathbf{H}_{\text{ex}})^{-1} \mathbf{H}_{\text{int}} \quad (1.22a)$$

By substituting Eq (1.21c) into Eq (1.21a) we find :

$$\mathbf{E}_{\text{corr}} = \mathbf{H}_{\text{int}}^{\leq} (\mathbf{E}_{\text{corr}} \mathbf{I} - \mathbf{H}_{\text{ex}})^{-1} \mathbf{H}_{\text{int}} \quad (1.22b)$$

From this equation \mathbf{E}_{corr} may be calculated without the need of calculating the coefficients d_k . The result is equivalent to the lowest eigenvalue of the CI matrix. However, since Eq (1.22b) is an implicit equation for \mathbf{E}_{corr} , an iterative method is needed to solve it (as is true for the diagonalisation method used for solving Eq (1.19)). A useful zero-order approximation is given by :

$$\mathbf{E}_{\text{corr}}^0 = -\mathbf{H}_{\text{int}}^{\leq} \mathbf{H}_{\text{ex}}^{-1} \mathbf{H}_{\text{int}} \quad (1.23)$$

It will be shown below that this is the most simple way to get a size consistent result. However, $\mathbf{E}_{\text{corr}}^0$ is not very accurate : the total energy $E_0 + \mathbf{E}_{\text{corr}}^0$ is not variational and in fact it is lower than the FCI result for our model system. In practice the solution to Eq (1.23) is in general comparable in accuracy to the converged solution to Eq (1.22), i.e. the CI result. If there are small diagonal elements in \mathbf{H}_{ex} (near-degeneracy with the HF function), $\mathbf{E}_{\text{corr}}^0$ is even worse than $\mathbf{E}_{\text{corr}}^{\text{DCI}}$, since $\mathbf{E}_{\text{corr}}^0$ is very sensitive to near-degeneracies due to the inversion of \mathbf{H}_{ex} in Eq (1.23).

§6 Size consistency corrections to CI calculations

From the above relations it is possible to derive simple formulas relating the FCI result to the DCI result for a system of n identical 2-electron systems without interaction. These may be used to correct the DCI result for the size consistency error. In practice these formulas appear also to be useful for SDCI calculations for electron pairs with interactions, that is, for a general molecular calculation.

We now consider a model system of n H_2 molecules, using the minimal AO basis. Using intermediate normalisation we have for the DCI function :

$$|\Psi_{\text{DCI}}\rangle = |\Phi_0\rangle + \sum_i^n d_i |D_i\rangle = |\Phi_0\rangle + d \sum_i^n |D_i\rangle \quad (1.24a)$$

where

$$|\Phi_0\rangle = \Phi_{01} \Phi_{02} \dots \Phi_{0i} \dots \Phi_{0n} \quad (1.24b)$$

and $|D_i\rangle$ is the double excitation on the i^{th} H_2 molecule :

$$D_i = \Phi_{01} \Phi_{02} \dots \Phi_{di} \dots \Phi_{0n} \quad (1.24c)$$

The energy is calculated by using Eq (1.22b) with :

$$\begin{aligned} (\mathbf{H}_{\text{int}})_i &= \langle D_i | \mathbf{H} | \Phi_0 \rangle = K && \text{for all } i \\ (\mathbf{H}_{\text{ex}})_{ij} &= (\langle \Phi_i | \mathbf{H} | \Phi_i \rangle - E_0) \delta_{ij} = \Delta \delta_{ij} && \text{for all } i, j \end{aligned} \quad (1.25)$$

Size Consistency in CI

Note that all off-diagonal elements of \mathbf{H}_{ex} vanish, since Φ_i and Φ_j ($i \neq j$) differ in more than 2 orbitals. Since the subsystems are identical, all interactions K have the same value, just as the excitation energies Δ .

From Eqs (1.22b) and (1.25) we have :

$$E_{\text{corr}}^{\text{DCI}} = \frac{nK^2}{E_{\text{corr}}^{\text{DCI}} - \Delta} \quad (1.26)$$

We thus have a quadratic equation for $E_{\text{corr}}^{\text{DCI}}$. The solution is :

$$E_{\text{corr}}^{\text{DCI}} = \frac{\Delta[1 - \sqrt{1 + nC}]}{2} \quad (1.27a)$$

where

$$C = \frac{4K^2}{\Delta^2} \quad (1.27b)$$

The coefficient d in Eq (1.24a) is then given by Eq (1.21a) :

$$d = \frac{E_{\text{corr}}^{\text{DCI}}}{nK} = \frac{1 - \sqrt{1 + nC}}{n\sqrt{C}} \quad (1.28)$$

Eq (1.27a) again shows that the DCI method is not size consistent, since $E_{\text{corr}}^{\text{DCI}}$ does not depend linearly on n . Eq (1.28) shows that d also depends on n .

The FCI function, on the other hand, may be written as :

$$\Psi_F = \Psi_{f_1} \dots \Psi_{f_i} \dots \Psi_{f_n} = [c_g \Phi_{g_1} + c_u \Phi_{u_1}] \dots [c_g \Phi_{g_i} + c_u \Phi_{u_i}] \dots [c_g \Phi_{g_n} + c_u \Phi_{u_n}] \quad (1.29)$$

and therefore the coefficient of the double excitations is given by :

$$d_{\text{FCI}} = \frac{c_u c_g^{n-1}}{c_g^n} = \frac{c_u}{c_g} \quad (1.30)$$

and this result does not depend on n . For $n = 1$ the DCI and the FCI calculations give the same result. Since the FCI energy depends linearly on n , we have, using Eq (1.27a) for $n = 1$:

$$E_{\text{corr}}^{\text{FCI}} = nE_{\text{corr}}^{\text{FCI}}(n = 1) = \frac{n\Delta[1 - \sqrt{1 + C}]}{2} \quad (1.31)$$

The relative size consistency correction thus follows from :

$$\frac{\Delta E_{\text{corr}}^{\text{DCI}}}{E_{\text{corr}}^{\text{DCI}}} = \frac{E_{\text{corr}}^{\text{FCI}} - E_{\text{corr}}^{\text{DCI}}}{E_{\text{corr}}^{\text{DCI}}} = \frac{n[1 - \sqrt{1 + C}]}{1 - \sqrt{1 + nC}} - 1 \quad (1.32)$$

The Pople correction.

Eq (1.32) may be used in order to find a generally useful size consistency correction for SDCI calculations. Using Eq (1.28), Eq (1.32) may be rewritten into a form, which may be applied to any SDCI calculation. In order to do this we relate the correction to the number of pairs, the CI correlation energy and the CI coefficients, as follows. First we express C in d , using Eq (1.28) :

Size Consistency in CI

$$C = \frac{4d^2}{(d^2n - 1)^2} \quad (1.33)$$

Now we revert to the usual normalisation :

$$|\Psi_{\text{DCI}}\rangle = \cos\theta |\Phi_0\rangle + \frac{\sin\theta}{\sqrt{n}} \sum_i |D_i\rangle \quad (1.34)$$

Exercise 1.3

Derive Eq (1.33) from Eq (1.28)

Exercise 1.4

Show that the function Eq (1.34) is normalised.

By comparing Eq (1.34) to Eq (1.18) we see that

$$d = \frac{1}{\sqrt{n}} \text{tg}\theta \quad (1.35)$$

By substituting this into Eq (1.33) and then into Eq (1.32) we find:

$$\frac{\Delta E_{\text{corr}}^{\text{Pople}}}{E_{\text{corr}}^{\text{DCI}}} = \frac{\sqrt{n^2 + n \text{tg}^2 2\theta} - n}{\sec 2\theta - 1} - 1 \quad (1.36)$$

Finally we assume that the single excitations may be treated in the same way. In order to make Eq (1.36) applicable to open shell systems n may be replaced by $N/2$ (N = the number of electrons), yielding :

$$\frac{\Delta E_{\text{corr}}^{\text{Pople}}}{E_{\text{corr}}^{\text{SDCI}}} = \frac{\sqrt{N^2 + 2N \text{tg}^2 2\theta} - N}{2(\sec 2\theta - 1)} - 1 \quad (1.37a)$$

where (because of Eq (1.34)) :

$$\cos\theta = c_0 \quad (1.37b)$$

This is the Pople correction.

In practice Eq (1.37) yields useful results for medium-sized molecules without near-degeneracy effects. It works better for closed shell systems than for open shells, particularly if a spin-adapted algorithm is used, since in the latter case the single excitations may become important (spin recoupling effects).

Size Consistency in CI

Exercise 1.5

Show that the Pople correction vanishes for $N = 2$, as it should.

Exercise 1.6

Why is the Pople correction not exact for an open shell system like $H_2 + H^*$?

The Davidson correction.

Another approximate method was introduced by Davidson in a rather intuitive way. It may also be obtained by substituting the power expansion for $\sqrt{1+x}$ into Eq (1.31) and assuming that $K \ll \Delta$, i.e. $C \ll 1$ and $\text{tg}2\theta = 2\sin\theta$. Again n must not be too large and Δ should not be too small (no near-degeneracies). The result is :

$$\Delta E_{\text{corr}}^{\text{Davidson}} = (1 - c_0^2)E_{\text{corr}} \quad (1.38)$$

This is the Davidson correction.

Note that this formula is not exact for $N = 2$, where the size consistency error should vanish.